CHAPTER

Domain 7: Security operations 7

CHAPTER OUTLINE

Eleventh Hour CISSP®. http://dx.doi.org/10.1016/B978-0-12-811248-9.00007-3

INTRODUCTION

Security operations is concerned with threats to a production-operating environment. Threat agents can be internal or external actors, and operations security must account for both of these threat sources in order to be effective. Security operations is about people, data, media, and hardware, as well as the threats associated with each of these in a production environment.

ADMINISTRATIVE SECURITY

All organizations contain people, data, and the means for people to use the data. A fundamental aspect of operations security is ensuring that controls are in place to inhibit people either inadvertently or intentionally compromising the confidentiality, integrity, or availability of data or the systems and media holding that data. Administrative security provides the means to control people's operational access to data.

ADMINISTRATIVE PERSONNEL CONTROLS

Administrative personnel controls represent important operations security concepts that should be mastered by the CISSP candidate. These are fundamental concepts within information security that permeate multiple domains.

Least privilege or minimum necessary access

One of the most important concepts in all of information security is that of the *principle of least privilege*. The principle of least privilege dictates that persons have no more than the access that is strictly required for the performance of their duties. The principle of least privilege may also be referred to as the principle of minimum necessary access. Regardless of name, adherence to this principle is a fundamental tenet of security and should serve as a starting point for administrative security controls.

Need to know

In organizations with extremely sensitive information that leverage mandatory access control (MAC), a basic determination of access is enforced by the system. The access determination is based upon clearance levels of subjects and classification levels of objects. Though the vetting process for someone accessing highly sensitive information is stringent, clearance level alone is insufficient when dealing with the most sensitive of information. An extension to the principle of least privilege in MAC environments is the concept of compartmentalization.

Compartmentalization, a method for enforcing *need to know*, goes beyond the mere reliance upon clearance level and necessitates simply that someone requires access to information. Compartmentalization is best understood by considering a highly sensitive military operation; while there may be a large number of individuals, some of whom might be of high rank, only a subset will "need to know" specific information. The others have no "need to know," and therefore will not be granted access.

Separation of duties

Separation of duties prescribes that multiple people are required to complete critical or sensitive transactions. The goal of separation of duties is to ensure that in order for someone to abuse their access to sensitive data or transactions, they must convince another party to act in concert. *Collusion* is the term used for the two parties conspiring to undermine the security of the transaction. The classic action movie example of separation of duties involves two keys, a nuclear sub, and a rogue captain.

Rotation of duties/job rotation

Rotation of duties, also known as job rotation or rotation of responsibilities, provides an organization with a means to help mitigate the risk associated with any one individual having too many privileges. Rotation of duties simply requires that one person does not perform critical functions or responsibilities for an extended period of time. There are multiple issues that rotation of duties can help to begin to address. One issue addressed by job rotation is the "hit by a bus" scenario. Imagine, morbid as it is, that one individual in the organization is hit by a bus on his/her way to work. If the operational impact of the loss of an individual would be too great, then perhaps one way to assuage this impact would be to ensure that there is additional depth of coverage for this individual's responsibilities.

Mandatory leave/forced vacation

An additional operational control that is closely related to rotation of duties is that of *mandatory leave*, also known as forced vacation. Though there are various justifications for requiring employees to be away from work, the primary security considerations are similar to that addressed by rotation of duties: reducing or detecting personnel single points of failure, and detecting and deterring fraud.

Nondisclosure agreement

A *nondisclosure agreement* (NDA) is a work-related contractual agreement ensuring that, prior to being given access to sensitive information or data, an individual or organization appreciates their legal responsibility to maintain the confidentiality of that sensitive information. Job candidates, consultants, or contractors often sign NDAs before they are hired. NDAs are largely a directive control.

Background checks

Background checks (also known as background investigations or preemployment screening) are an additional administrative control commonly employed by many organizations. The majority of background investigations are performed as part of a preemployment screening process. Some organizations perform cursory background investigations that include a criminal record check. Others perform more in-depth checks, such as verifying employment history, obtaining credit reports, and, in some cases, requiring the submission of a drug screening.

FORENSICS

Digital forensics provides a formal approach to dealing with investigations and evidence with special consideration of the legal aspects of this process. The forensic process must preserve the "crime scene" and the evidence in order to prevent the unintentional violation of the integrity of either the data or the data's environment. A primary goal of forensics is to prevent unintentional modification of the system. *Live forensics* includes taking a bit-by-bit image or *binary image* of physical memory, gathering details about running processes, and gathering network connection data.

FORENSIC MEDIA ANALYSIS

In addition to the valuable data gathered during the live forensic capture, the main source of forensic data typically comes from binary images of secondary storage and portable storage devices such as hard disk drives, USB flash drives, CDs, DVDs, and possibly associated cellular (mobile) phones and mp3 players.

FAST FACTS

Here are the four basic types of disk-based forensic data:

- *Allocated space*: portions of a disk partition that are marked as actively containing data.
- *Unallocated space*: portions of a disk partition that do not contain active data. This includes portions that have never been allocated, as well as previously allocated portions that have been marked unallocated. If a file is deleted, the portions of the disk that held the deleted file are marked as unallocated and made available for use.
- *Slack space*: data is stored in specific-sized chunks known as clusters, which are sometimes referred to as sectors or blocks. A cluster is the minimum size that can be allocated by a file system. If a particular file, or final portion of a file, does not require the use of the entire cluster, then some extra space will exist within the cluster. This leftover space is known as slack space; it may contain old data, or it can be used intentionally by attackers to hide information.
- *"Bad" blocks/clusters/sectors*: hard disks routinely end up with sectors that cannot be read due to some physical defect. The sectors marked as bad will be ignored by the operating system since no data could be read in those defective portions. Attackers could intentionally mark sectors or clusters as being bad in order to hide data within this portion of the disk.

NETWORK FORENSICS

Network forensics is the study of data in motion, with a special focus on gathering evidence via a process that will support admission into a court of law. This means the integrity of the data is paramount, as is the legality of the collection process. Network forensics is closely related to network intrusion detection; the difference is the former focuses on legalities, while the later focuses on operations.

EMBEDDED DEVICE FORENSICS

One of the greatest challenges facing the field of digital forensics is the proliferation of consumer-grade electronic hardware and embedded devices. While forensic investigators have had decades to understand and develop tools and techniques to analyze magnetic disks, newer technologies such as solid-state drives lack both forensic understanding and forensic tools capable of analysis.

ELECTRONIC DISCOVERY (eDISCOVERY)

Electronic discovery, or eDiscovery, pertains to legal counsel gaining access to pertinent electronic information during the pretrial discovery phase of civil legal proceedings. The general purpose of discovery is to gather potential evidence that will allow for building a case. Electronic discovery differs from traditional discovery simply in that eDiscovery seeks ESI, or electronically stored information, which is typically acquired via a forensic investigation. While the difference between traditional discovery and eDiscovery might seem miniscule, given the potentially vast quantities of electronic data stored by organizations, eDiscovery can become logistically and financially cumbersome.

Some of the challenges associated with eDiscovery stem from the seemingly innocuous backup policies of organizations. While long-term storage of computer information has generally been thought to be a sound practice, this data is discoverable. Discovery does not take into account whether ESI is conveniently accessible or transferable.

Appropriate data retention policies, in addition to software and systems designed to facilitate eDiscovery, can greatly reduce the burden on the organization when required to provide ESI for discovery. When considering data retention policies, consider not only how long information should be kept, but also how long the information needs to be accessible to the organization. Any data for which there is no longer a need should be appropriately purged according to the data retention policy.

INCIDENT RESPONSE MANAGEMENT

All organizations will experience security incidents. Because of the certainty of security incidents eventually impacting organizations, there is a great need to be equipped with a regimented and tested methodology for identifying and responding to these incidents.

METHODOLOGY

Different books and organizations may use different terms and phases associated with the incident response process; this section will mirror the terms associated with the examination. Many incident-handling methodologies treat containment, eradication, and recovery as three distinct steps, as we will in this book. Other names for each step are sometimes used; the current exam lists a seven-step lifecycle but curiously omits the first step in most incident handling methodologies: preparation. Perhaps preparation is implied, like the identification portion of AAA systems. We will therefore cover eight steps, mapped to the current exam:

1. Preparation
2. Detection (identification)
3. Response (containment)
4. Mitigation (eradication)
5. Reporting
6. Recovery
7. Remediation
8. Lessons learned (postincident activity, postmortem, or reporting)

Preparation

The preparation phase includes steps taken before an incident occurs. These include training, writing incident response policies and procedures, and providing tools such as laptops with sniffing software, crossover cables, original OS media, removable

	Action	Completed
	Detection and Analysis	
1.	Determine whether an incident has occurred	
1.1	Analyze the precursors and indicators	
1.2	Look for correlating information	
1.3	Perform research (eg, search engines, knowledge base)	
1.4	As soon as the handler believes an incident has occurred, begin documenting the investigation and gathering evidence	
2.	Prioritize handling the incident based on the relevant factors (functional impact, information impact, recoverability effort, etc.)	
3.	Report the incident to the appropriate internal personnel and external organizations	
	Containment, Eradication, and Recovery	
4.	Acquire, preserve, secure, and document evidence	
5.	Contain the incident	
6.	Eradicate the incident	
6.1	Identify and mitigate all vulnerabilities that were exploited	
6.2	Remove malware, inappropriate materials, and other components	
6.3	If more affected hosts are discovered (eg, new malware infections), repeat the Detection and Analysis steps (1.1, 1.2) to identify all other affected hosts, then contain (5) and eradicate (6) the incident for them	
7.	Recover from the incident	
7.1	Return affected systems to an operationally ready state	
7.2	Confirm that the affected systems are functioning normally	
7.3	If necessary, implement additional monitoring to look for future related activity	
	Post-Incident Activity	
8.	Create a follow-up report	
9.	Hold a lessons learned meeting (mandatory for major incidents, optional otherwise)	

FIG. 7.1

Incident handling checklist.[1]

drives, etc. Preparation should include anything that may be required to handle an incident or that will make incident response faster and more effective. One preparation step is preparing an incident handling checklist. Fig. 7.1 is an incident handling checklist from NIST Special Publication 800-61r2.

Detection (identification)

One of the most important steps in the incident response process is the *detection phase*. Detection, also called identification, is the phase in which events are analyzed in order to determine whether these events might comprise a security incident. Without strong detective capabilities built into the information systems, the organization has little hope of being able to effectively respond to information security incidents in a timely fashion.

Response (containment)

The *response phase*, or containment, of incident response is the point at which the incident response team begins interacting with affected systems and attempts to keep further damage from occurring as a result of the incident. Responses might include

taking a system off the network, isolating traffic, powering off the system, or other items to control both the scope and severity of the incident. This phase is also typically where a binary (bit-by-bit) forensic backup is made of systems involved in the incident. An important trend to understand is that most organizations will now capture volatile data before pulling the power plug on a system.

Mitigation (eradication)

The *mitigation phase*, or eradication, involves the process of understanding the cause of the incident so that the system can be reliably cleaned and ultimately restored to operational status later in the recovery phase. In order for an organization to recover from an incident, the cause of the incident must be determined. The cause must be known so that the systems in question can be returned to a known good state without significant risk of the compromise persisting or reoccurring. A common occurrence is for organizations to remove the most obvious piece of malware affecting a system and think that is sufficient; when in reality, the obvious malware may only be a symptom and the cause may still be undiscovered.

Once the cause and symptoms are determined, the system needs to be restored to a good state and should not be vulnerable to further impact. This will typically involve either rebuilding the system from scratch or restoring from a known good backup.

Reporting

The reporting phase of incident handling occurs throughout the process, beginning with detection. Reporting must begin immediately upon detection of malicious activity. Reporting contains two primary areas of focus: technical and nontechnical reporting. The incident handling teams must report the technical details of the incident as they begin the incident handling process, while maintaining sufficient bandwidth to also notify management of serious incidents. A common mistake is forgoing the latter while focusing on the technical details of the incident itself, but this is a mistake. Nontechnical stake holders including business and mission owners must be notified immediately of any serious incident and kept up to date as the incident-handing process progresses.

Recovery

The *recovery phase* involves cautiously restoring the system or systems to operational status. Typically, the business unit responsible for the system will dictate when the system will go back online. Remember to be cognizant of the possibility that the infection, attacker, or other threat agent might have persisted through the eradication phase. For this reason, close monitoring of the system after it returns to production is necessary. Further, to make the security monitoring of this system easier, strong preference is given to the restoration of operations occurring during off-peak production hours.

Remediation

Remediation steps occur during the mitigation phase, where vulnerabilities within the impacted system or systems are mitigated. Remediation continues after that phase and becomes broader. For example, if the root-cause analysis determines that a password was stolen and reused, local mitigation steps could include changing the

compromised password and placing the system back online. Broader remediation steps could include requiring dual-factor authentication for all systems accessing sensitive data. We will discuss root-cause analysis shortly.

Lessons learned

The goal of this phase is to provide a final report on the incident, which will be delivered to management. Important considerations for this phase should include detailing ways in which the compromise could have been identified sooner, how the response could have been quicker or more effective, which organizational shortcomings might have contributed to the incident, and what other elements might have room for improvement. Feedback from this phase feeds directly into continued preparation, where the lessons learned are applied to improving preparation for the handling of future incidents.

ROOT-CAUSE ANALYSIS

To effectively manage security incidents, root-cause analysis must be performed. Root-cause analysis attempts to determine the underlying weakness or vulnerability that allowed the incident to be realized. Without successful root-cause analysis, the victim organization could recover systems in a way that still includes the particular weaknesses exploited by the adversary causing the incident. In addition to potentially recovering systems with exploitable flaws, another possibility includes reconstituting systems from backups or snapshots that have already been compromised.

OPERATIONAL PREVENTIVE AND DETECTIVE CONTROLS

Many preventive and detective controls require higher operational support and are a focus of daily operations security. For example, routers and switches tend to have comparatively low operational expenses (OPEX). Other controls, such as NIDS and NIPS, antivirus, and application whitelisting have comparatively higher OPEX and are a focus in this domain.

INTRUSION DETECTION SYSTEMS AND INTRUSION PREVENTION SYSTEMS

An intrusion detection system (IDS) detects malicious actions, including violations of policy. An intrusion prevention system (IPS) also prevents malicious actions. There are two basic types of IDSs and IPSs: network based and host based.

IDS and IPS event types

There are four types of IDS events: true positive, true negative, false positive, and false negative. We will use two streams of traffic, a worm and a user surfing the Web, to illustrate these events.

- True positive: A worm is spreading on a trusted network; NIDS alerts
- True negative: User surfs the Web to an allowed site; NIDS is silent
- False positive: User surfs the Web to an allowed site; NIDS alerts
- False negative: A worm is spreading on a trusted network; NIDS is silent

The goal is to have only true positives and true negatives, but most IDSs have false positives and false negatives as well. False positives waste time and resources, as monitoring staff spends time investigating nonmalicious events. A false negative is arguably the worst-case scenario because malicious network traffic is neither prevented nor detected.

NIDS and NIPS

A network-based intrusion detection system (NIDS) detects malicious traffic on a network. NIDS usually require promiscuous network access in order to analyze all traffic, including all unicast traffic. NIDS are passive devices that do not interfere with the traffic they monitor; Fig. 7.2 shows a typical NIDS architecture. The NIDS sniffs the internal interface of the firewall in read-only mode and sends alerts to a NIDS Management server via a different (ie, read/write) network interface.

The difference between a NIDS and a *NIPS* is that the NIPS alters the flow of network traffic. There are two types of NIPS: active response and inline. Architecturally, an active response NIPS is like the NIDS in Fig. 7.2; the difference is that the monitoring interface is read/write. The active response NIPS may "shoot down" malicious traffic via a variety of methods, including forging TCP RST segments to source or destination (or both), or sending ICMP port, host, or network unreachable to source.

An inline NIPS is "in line" with traffic, acting as a Layer 3–7 firewall by passing or allowing traffic, as shown in Fig. 7.3.

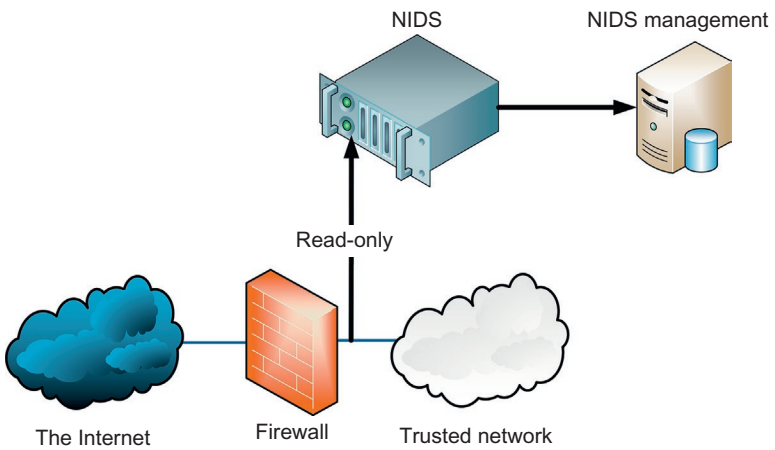

FIG. 7.2

NIDS architecture.

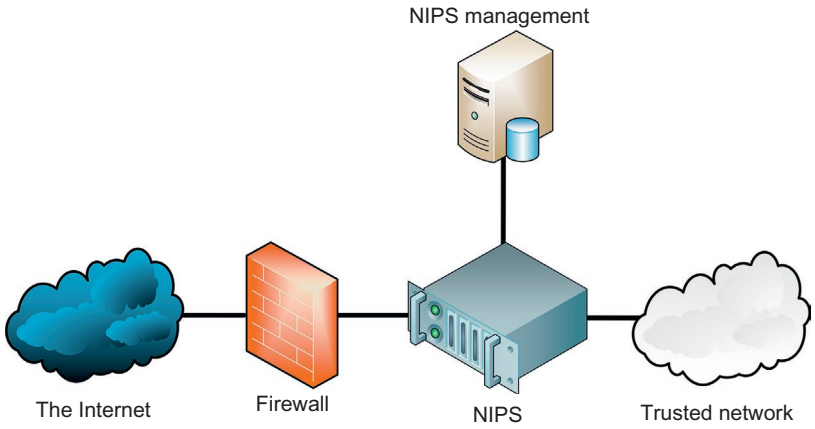

FIG. 7.3

Inline NIPS architecture.

Note that a NIPS provides defense-in-depth protection in addition to a firewall; it is not typically used as a replacement. Also, a false positive by a NIPS is more damaging than one by a NIDS because legitimate traffic is denied, which may cause production problems. A NIPS usually has a smaller set of rules compared to a NIDS for this reason, and only the most trustworthy rules are used. A NIPS is not a replacement for a NIDS; many networks use both a NIDS and a NIPS.

HIDS and HIPS

Host-based intrusion detection systems (HIDS) and host-based intrusion prevention systems (HIPS) are host-based cousins to NIDS and NIPS. They process information within the host and may process network traffic as it enters the host, but the exam's focus is usually on files and processes.

SECURITY INFORMATION AND EVENT MANAGEMENT

Correlation of security-relevant data is the primary utility provided by Security Information and Event Management (SIEM). The goal of data correlation is to better understand the context so as to arrive at a greater understanding of risk within the organization due to activities that are noted across various security platforms. While SIEMs typically come with some built-in alerts that look for particular correlated data, custom correlation rules are typically created to augment the built-in capabilities.

DATA LOSS PREVENTION

As prominent and high-volume data breaches continue unabated, the desire for solutions designed to address data loss has grown. Data loss prevention (DLP) is a class of solutions that are tasked specifically with trying to detect or preferably prevent data

from leaving an organization in an unauthorized manner. The approaches to DLP vary greatly. One common approach employs network-oriented tools that attempt to detect and/or prevent sensitive data being exfiltrated in cleartext. This approach does little to address the potential for data exfiltration over an encrypted channel. Dealing with the potential for encrypted exfiltration typically requires endpoint solutions to provide visibility prior to encryption.

ENDPOINT SECURITY

Because endpoints are the targets of attacks, preventive and detective capabilities on the endpoints themselves provide a layer beyond network-centric security devices. Modern endpoint security suites often encompass a variety of products beyond simple antivirus software. These suites can increase the depth of security countermeasures well beyond the gateway or network perimeter.

An additional benefit offered by endpoint security products is their ability to provide preventive and detective control even when communications are encrypted all the way to the endpoint in question. Typical challenges associated with endpoint security are associated with volume considerations; vast number of products/systems must be managed, while significant amounts of data must be analyzed and potentially retained.

Antivirus

The most commonly deployed endpoint security product is antivirus software. Antivirus is one of many layers of endpoint defense-in-depth security. Although antivirus vendors often employ heuristic or statistical methods for malware detection, the predominant means of detecting malware is still signature based.

Application whitelisting

Application whitelisting is a more recent addition to endpoint security suites. The primary focus of application whitelisting is to determine in advance which binaries are considered safe to execute on a given system. Once this baseline has been established, any binary attempting to run that is not on the list of "known-good" binaries is prevented from doing so. A weakness of this approach is when a "known-good" binary is exploited by an attacker and used maliciously.

Removable media controls

The need for better control of removable media has been felt on two fronts in particular. First, malware-infected removable media inserted into an organization's computers has been a method for compromising otherwise reasonably secure organizations. Second, the volume of storage that can be contained in something the size of a fingernail is astoundingly large and has been used to surreptitiously exfiltrate sensitive data.

Disk encryption

Another endpoint security product found with increasing regularity is disk encryption software.

Full disk encryption, also called whole disk encryption, encrypts an entire disk. This is superior to partially encrypted solutions, such as encrypted volumes, directories, folders, or files. The problem with the latter approach is the risk of leaving sensitive data on an unencrypted area of the disk.

ASSET MANAGEMENT

A holistic approach to operational information security requires organizations to focus on systems as well as the people, data, and media. Systems security is another vital component to operational security, and there are specific controls that can greatly help system security throughout the system's lifecycle.

CONFIGURATION MANAGEMENT

Basic *configuration management* practices associated with system security will involve tasks such as disabling unnecessary services; removing extraneous programs; enabling security capabilities such as firewalls, antivirus, and intrusion detection or prevention systems; and configuring security and audit logs.

Baselining

Security *baselining* is the process of capturing a snapshot of the current system security configuration. Establishing an easy means for capturing the current system security configuration can be extremely helpful in responding to a potential security incident.

Vulnerability management

Vulnerability scanning is a way to discover poor configurations and missing patches in an environment. The term *vulnerability management* is used rather than just vulnerability scanning in order to emphasize the need for management of the vulnerability information. The remediation or mitigation of vulnerabilities should be prioritized based on both risk to the organization and ease of remediation procedures.

Zero-day vulnerabilities and zero-day exploits

A zero-day vulnerability is a vulnerability that is known before the existence of a patch. *Zero-day vulnerabilities*, also commonly written 0-day, are becoming increasingly important as attackers are becoming more skilled in discovery and disclosure of zero-day vulnerabilities is being monetized. A *zero-day exploit*, rather than vulnerability, refers to the existence of exploit code for a vulnerability that has yet to be patched.

CHANGE MANAGEMENT

In order to maintain consistent and known operational security, a regimented *change management* or change control process needs to be followed. The purpose of the change control process is to understand, communicate, and document any changes

with the primary goal of being able to understand, control, and avoid direct or indirect negative impact that the change might impose.

FAST FACTS

Because of the variability of the change management process, specifically named phases have not been offered in this section. However, the general flow of the change management process includes:

- Identifying a change
- Proposing a change
- Assessing the risk associated with the change
- Testing the change
- Scheduling the change
- Notifying impacted parties of the change
- Implementing the change
- Reporting results of the change implementation

All changes must be closely tracked and auditable; a detailed change record should be kept. Some changes can destabilize systems or cause other problems; change management auditing allows operations staff to investigate recent changes in the event of an outage or problem. Audit records also allow auditors to verify that change management policies and procedures have been followed.

CONTINUITY OF OPERATIONS

Continuity of operations is principally concerned with the availability portion of the confidentiality, integrity, and availability triad.

SERVICE LEVEL AGREEMENTS

A *service level agreement* (SLA) stipulates all expectations regarding the behavior of the department or organization that is responsible for providing services and the quality of those services. SLAs will often dictate what is considered acceptable regarding things such as bandwidth, time to delivery, response times, etc.

FAULT TOLERANCE

In order for systems and solutions within an organization to be able to continually provide operational availability, they must be implemented with fault tolerance in mind. Availability is not solely focused on system uptime requirements; it requires that data be accessible in a timely fashion as well.

Redundant array of inexpensive disks

Even if only one full backup tape is needed for recovery of a system due to a hard disk failure, the time to recover a large amount of data can easily exceed the recovery

time dictated by the organization. The goal of a *redundant array of inexpensive disks (RAID)* is to help mitigate the risk associated with hard disk failures. There are various RAID levels that consist of different approaches to disk array configurations.

FAST FACTS

Three critical RAID terms are mirroring, striping, and parity.

- *Mirroring* achieves full data redundancy by writing the same data to multiple hard disks.
- *Striping* focuses on increasing read and write performance by spreading data across multiple hard disks. Writes can be performed in parallel across multiple disks rather than serially on one disk. This parallelization increases performance and does not contribute to data redundancy.
- *Parity* achieves data redundancy without incurring the same degree of cost as that of mirroring in terms of disk usage and write performance.

RAID 0: Striped set

RAID 0 employs striping to increase the performance of read and writes. Striping offers no data redundancy, so RAID 0 is a poor choice if recovery of data is critical. Fig. 7.4 shows RAID 0.

RAID 1: Mirrored set

RAID 1 creates/writes an exact duplicate of all data to an additional disk. Fig. 7.5 shows RAID 1.

RAID 2: Hamming code

RAID 2 is a legacy technology that requires either 14 or 39 hard disks and a specially designed hardware controller, which makes RAID 2 cost prohibitive. RAID 2 stripes at the bit level.

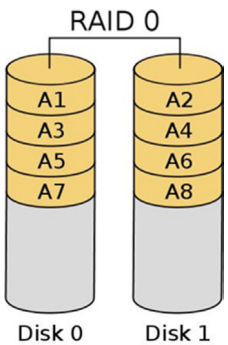

FIG. 7.4

RAID 0—striped set.

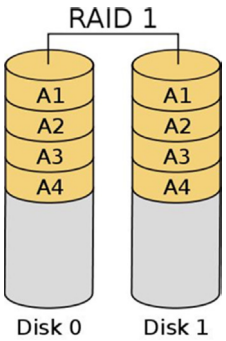

FIG. 7.5

RAID 1—mirrored set.

EXAM WARNING

While the ability to quickly recover from a disk failure is a goal of RAID, there are configurations that do not have reliability as a capability. For the exam, understand that not all RAID configurations provide additional reliability.

RAID 3: Striped set with dedicated parity (byte level)

Striping is desirable due to the performance gains associated with spreading data across multiple disks. However, striping alone is not as desirable due to the lack of redundancy. With *RAID 3*, data at the byte level is striped across multiple disks, but an additional disk is leveraged for storage of parity information, which is used for recovery in the event of a failure.

RAID 4: Striped set with dedicated parity (block level)

RAID 4 provides the same functionality as RAID 3, but stripes data at the block level rather than byte level. Like RAID 3, RAID 4 employs a dedicated parity drive rather than having parity data distributed amongst all disks, as in RAID 5.

RAID 5: Striped set with distributed parity

One of the most popular RAID configurations is that of *RAID 5*, striped set with distributed parity. Like RAIDs 3 and 4, RAID 5 writes parity information that is used for recovery purposes. RAID 5 writes at the block level, like RAID 4. However, unlike RAIDs 3 and 4, which require a dedicated disk for parity information, RAID 5 distributes the parity information across multiple disks. One of the reasons for RAID 5's popularity is that the disk cost for redundancy is potentially lower than that of a mirrored set, while at the same time gaining performance improvements associated with RAID 0. RAID 5 allows for data recovery in the event that any one disk fails. Fig. 7.6 shows RAID 5.

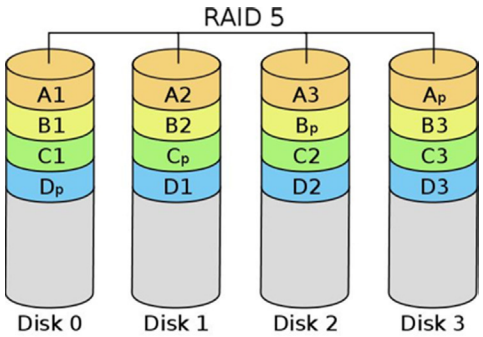

FIG. 7.6

RAID 5—striped set with distributed parity.

RAID 6: Striped set with dual-distributed parity

While RAID 5 accommodates the loss of any one drive in the array, *RAID 6* can allow for the failure of two drives and still function. This redundancy is achieved by writing the same parity information to two different disks.

RAID 1+0 or RAID 10

RAID 1+0 or RAID 10 is an example of what is known as nested RAID or multi-RAID, which simply means that one standard RAID level is encapsulated within another. With RAID 10, which is also commonly written as RAID 1+0 to explicitly indicate the nesting, the configuration is that of a striped set of mirrors.

CRUNCH TIME

Table 7.1 provides a brief description of the various RAID levels that are most commonly used.

Table 7.1 RAID Levels

RAID Level	Description
RAID 0	Block-level striped set
RAID 1	Mirrored set
RAID 3	Byte-level striping with dedicated parity
RAID 4	Block-level striping with dedicated parity
RAID 5	Block-level striping with distributed parity
RAID 6	Block-level striping with double distributed parity

System redundancy

Redundant hardware and redundant systems

Many systems can provide internal hardware redundancy of components that are extremely prone to failure. The most common example of this built-in redundancy is systems or devices that have redundant onboard power in the event of a power supply failure. Sometimes systems simply have field replaceable modular versions of commonly failing components. Though physically replacing a power supply might increase downtime, having an inventory of spare modules to service all of the datacenter's servers would be less expensive than having all servers configured with an installed redundant power supply.

Redundant systems (ie, alternative systems) make entire systems available in case of failure of the primary system.

High availability clusters

A *high-availability cluster*, also called a *failover cluster*, uses multiple systems that are already installed, configured, and plugged in, so that if a failure causes one of the systems to fail, another can be seamlessly leveraged to maintain the availability of the service or application being provided.

Each member of an *active-active* HA cluster actively processes data in advance of a failure. This is commonly referred to as load balancing. Having systems in an active-active or load-balancing configuration is typically more costly than having the systems in an *active-passive* or hot standby configuration, in which the backup systems only begin processing when a failure is detected.

BCP AND DRP OVERVIEW AND PROCESS

The terms and concepts associated with Business Continuity and Disaster Recovery Planning are very often misunderstood. Clear understanding of what is meant by both Business Continuity and Disaster Recovery Planning, as well as what they entail, is critical for the CISSP candidate.

BUSINESS CONTINUITY PLANNING

Though many organizations will simply use the phrases *Business Continuity Planning* (BCP) or *Disaster Recovery Planning* (DRP) interchangeably, they are two distinct disciplines. Though both types of planning are essential to the effective management of disasters and other disruptive events, their goals are different. The overarching goal of BCP is to ensure that the business will continue to operate before, throughout, and after a disaster event is experienced. The focus of BCP is on the business as a whole, ensuring that those critical services or functions the business provides or performs can still be carried out both in the wake of a disruption and after the disruption has been weathered.

DISASTER RECOVERY PLANNING

The Disaster Recovery Plan (DRP) provides a short-term plan for dealing with specific IT-oriented disruptions. Mitigating a malware infection that shows risk of spreading to other systems is an example of a specific IT-oriented disruption that a DRP would address. The DRP focuses on efficiently attempting to mitigate the impact of a disaster by preparing the immediate response and recovery of critical IT systems. DRP is considered tactical rather than strategic and provides a means for immediate response to disasters.

RELATIONSHIP BETWEEN BCP AND DRP

The BCP is an umbrella plan that includes multiple specific plans, most importantly the DRP. DRP serves as a subset of the overall BCP, which would be doomed to fail if it did not contain a tactical method for immediately dealing with disruption of information systems. Fig. 7.7, taken from *NIST Special Publication 800-34*, provides a visual means for understanding the interrelatedness of BCP and DRP, as well as *Continuity of Operations Plan (COOP)*, *Occupant Emergency Plan (OEP)*, and others.

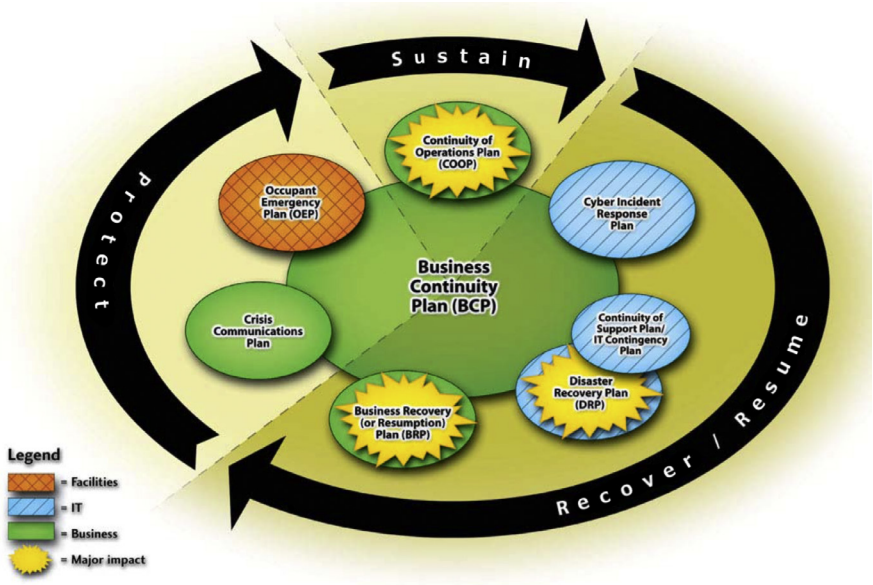

FIG. 7.7

BCP and related plans.[2]

DISASTERS OR DISRUPTIVE EVENTS

Given that BCP and DRP are created because of the potential of disasters impacting operations, it is vital that organizations understand the nature of disasters and disruptive events.

FAST FACTS

The three common ways of categorizing the causes for disasters are derived from whether the threat agent is natural, human, or environmental in nature.[2]

- Natural—This category includes threats such as earthquakes, hurricanes, tornadoes, floods, and some types of fires. Historically, natural disasters have provided some of the most devastating disasters to which an organization must respond.
- Human—The human category of threats represents the most common source of disasters. Human threats can be further classified by whether they constitute an intentional or unintentional threat.
- Environmental—Threats focused on information systems or datacenter environments; includes items such as power issues (blackout, brownout, surge, spike, etc.), system component or other equipment failures, and application or software flaws.

The analysis of threats and the determination of the associated likelihood of those threats are important parts of the BCP and DRP process. Table 7.2 provides a quick summary of some of the disaster events and what type of disaster they constitute.

FAST FACTS

Types of disruptive events include:

- Errors and omissions: typically considered the most common source of disruptive events. This type of threat is caused by humans who unintentionally serve as a source of harm.
- Natural disasters: include earthquakes, hurricanes, floods, tsunamis, etc.
- Electrical or power problems: loss of power may cause availability issues, as well as integrity issues due to corrupted data.
- Temperature and humidity failures: may damage equipment due to overheating, corrosion, or static electricity.
- Warfare, terrorism, and sabotage: threats can vary dramatically based on geographic location, industry, and brand value, as well as the interrelatedness with other high-value target organizations.
- Financially motivated attackers: attackers who seek to make money by attacking victim organizations, includes exfiltration of cardholder data, identity theft, pump-and-dump stock schemes, bogus antimalware tools, corporate espionage, and others.
- Personnel shortages: may be caused by strikes, pandemics, or transportation issues. A lack of staff may lead to operational disruption.

Table 7.2 Examples of Disruptive Events

Disruptive Event	Type
Earthquake/tornado/hurricane/etc.	Natural
Strike	Human (intentional)
Cyber terrorism	Human (intentional/technical)
Malware	Human (intentional/technical)
Denial of service	Human (intentional/technical)
Errors and omissions	Human (unintentional)
Electrical fire	Environmental
Equipment failure	Environmental

THE DISASTER RECOVERY PROCESS

Having discussed the importance of BCP and DRP as well as examples of threats that justify this degree of planning, we will now focus on the fundamental steps involved in recovering from a disaster.

Respond

In order to begin the disaster recovery process, there must be an initial response that begins the process of assessing the damage. Speed is essential during this initial assessment, which will determine if the event in question constitutes a disaster.

Activate team

If a disaster is declared, then the recovery team needs to be activated. Depending on the scope of the disaster, this communication could prove extremely difficult. The use of calling trees, which will be discussed in the "Call Trees" section later in this chapter, can help to facilitate this process to ensure that members can be activated as smoothly as possible.

Communicate

One of the most difficult aspects of disaster recovery is ensuring that consistent timely status updates are communicated back to the central team managing the response and recovery process. This communication often must occur out-of-band, meaning that the typical communication method of leveraging an office phone will quite often not be a viable option. In addition to communication of internal status regarding the recovery activities, the organization must be prepared to provide external communications, which involves disseminating details regarding the organization's recovery status with the public.

Assess

Though an initial assessment was carried out during the initial response portion of the disaster recovery process, a more detailed and thorough assessment will be performed by the disaster recovery team. The team will proceed to assessing the extent of the damage to determine the proper steps necessary to ensure the organization's ability to meet its mission.

Reconstitution

The primary goal of the reconstitution phase is to successfully recover critical business operations at either a primary or secondary site. If an alternate site is leveraged, adequate safety and security controls must be in place in order to maintain the expected degree of security the organization typically employs. The use of an alternate computing facility for recovery should not expose the organization to further security incidents. In addition to the recovery team's efforts in reconstituting critical business functions at an alternate location, a salvage team will be employed to begin the recovery process at the primary facility that experienced the disaster. Ultimately, the expectation is that unless it is wholly unwarranted given the circumstances, the primary site will be recovered and that the alternate facility's operations will "fail back" or be transferred again to the primary center of operations.

DEVELOPING A BCP/DRP

Developing BCP/DRP is vital for an organization's ability to respond and recover from an interruption in normal business functions or catastrophic event. In order to ensure that all planning has been considered, the BCP/DRP has a specific set of requirements to review and implement. Below are listed these high-level steps, according to NIST SP800-34, to achieving a sound, logical BCP/DRP. NIST SP800-34 is the National Institute of Standards and Technologies Contingency Planning Guide for Federal Information Systems.

- Project Initiation
- Scope of the Project
- Business Impact Analysis (BIA)
- Identify Preventive Controls
- Recovery Strategy
- Plan Design and Development
- Implementation, Training, and Testing
- BCP/DRP Maintenance[2]

PROJECT INITIATION

In order to develop the BCP/DRP, the scope of the project must be determined and agreed upon.

FAST FACTS

Project Initiation involves seven distinct milestones,[2] as listed below:

- *Develop the contingency planning policy statement*: A formal department or agency policy provides the authority and guidance necessary to develop an effective contingency plan.
- *Conduct the BIA*: The BIA helps identify and prioritize critical IT systems and components. A template for developing the BIA is also provided to assist the user.
- *Identify preventive controls*: Measures taken to reduce the effects of system disruptions can increase system availability and reduce contingency life-cycle costs.
- *Develop recovery strategies*: Thorough recovery strategies ensure that the system may be recovered quickly and effectively following a disruption.
- *Develop an IT contingency plan*: The contingency plan should contain detailed guidance and procedures for restoring a damaged system.
- *Plan testing, training, and exercises*: Testing the plan identifies planning gaps, whereas training prepares recovery personnel for plan activation; both activities improve plan effectiveness and overall agency preparedness.
- *Plan maintenance*: The plan should be a living document that is updated regularly to remain current with system enhancements.[2]

ASSESSING THE CRITICAL STATE

Assessing the critical state can be difficult because determining which pieces of the IT infrastructure are critical depends solely on the how it supports the users within the organization. For example, without consulting all of the users, a simple mapping program may not seem to be a critical asset for an organization. However, if there is a user group that drives trucks and makes deliveries for business purposes, this mapping software may be critical for them to schedule pickups and deliveries.

CONDUCT BIA

BIA is the formal method for determining how a disruption to the IT system(s) of an organization will impact the organization's requirements, processes, and interdependencies with respect to the business mission.[2] It is an analysis to identify and prioritize critical IT systems and components. It enables the BCP/DRP project manager to fully characterize the IT contingency requirements and priorities.[2] The objective is to correlate the IT system components with the critical service it supports. It also aims to quantify the consequence of a disruption to the system component and how that will affect the organization. The primary goal of the BIA is to determine the Maximum Tolerable Downtime (MTD) for a specific IT asset. This will directly impact what disaster recovery solution is chosen.

Identify critical assets

The critical asset list is a list of those IT assets that are deemed business-essential by the organization. These systems' DRP/BCP must have the best available recovery capabilities assigned to them.

Conduct BCP/DRP-focused risk assessment

The BCP/DRP-focused risk assessment determines what risks are inherent to which IT assets. A vulnerability analysis is also conducted for each IT system and major application. This is done because most traditional BCP/DRP evaluations focus on physical security threats, both natural and human.

Determine MTD

The primary goal of the BIA is to determine the *MTD*, which describes the total time a system can be inoperable before an organization is severely impacted. MTD is comprised of two metrics: the *Recovery Time Objective (RTO)*, and the *Work Recovery Time (WRT)* (see later).

Alternate terms for MTD

Depending on the business continuity framework that is used, other terms may be substituted for MTD. These include *Maximum Allowable Downtime*, Maximum Tolerable Outage, and Maximum Acceptable Outage.

Failure and recovery metrics

A number of metrics are used to quantify how frequently systems fail, how long a system may exist in a failed state, and the maximum time to recover from failure. These metrics include the *Recovery Point Objective (RPO)*, RTO, WRT, Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR), and *Minimum Operating Requirements (MOR)*.

Recovery point objective

The RPO is the amount of data loss or system inaccessibility (measured in time) that an organization can withstand. "If you perform weekly backups, someone made a decision that your company could tolerate the loss of a week's worth of data. If backups are performed on Saturday evenings and a system fails on Saturday afternoon, you have lost the entire week's worth of data. This is the RPO. In this case, the RPO is 1 week."[3]

The RPO represents the maximum acceptable amount of data/work loss for a given process because of a disaster or disruptive event.

Recovery time objective and work recovery time

The RTO describes the maximum time allowed to recover business or IT systems. RTO is also called the systems recovery time. This is one part of MTD; once the system is physically running, it must be configured.

CRUNCH TIME

WRT describes the time required to configure a recovered system. "Downtime consists of two elements, the systems recovery time and the WRT. Therefore, MTD = RTO + WRT."[3]

Mean time between failures

MTBF quantifies how long a new or repaired system will run before failing. It is typically generated by a component vendor and is largely applicable to hardware as opposed to applications and software.

Mean time to repair

The MTTR describes how long it will take to recover a specific failed system. It is the best estimate for reconstituting the IT system so that business continuity may occur.

Minimum operating requirements

MOR describe the minimum environmental and connectivity requirements in order to operate computer equipment. It is important to determine and document what the MOR is for each IT-critical asset because in the event of a disruptive event or disaster, proper analysis can be conducted quickly to determine if the IT assets will be able to function in the emergency environment.

IDENTIFY PREVENTIVE CONTROLS

Preventive controls can prevent disruptive events from having an impact. For example, as stated in Chapter 3, HVAC systems are designed to prevent computer equipment from overheating and failing.

DID YOU KNOW?

The BIA will identify some risks that may be mitigated immediately. This is another advantage of performing BCP/DRP, including the BIA: it improves your security, even if no disaster occurs.

RECOVERY STRATEGY

Once the BIA is complete, the BCP team knows the MTD. This metric, as well as others including the RPO and RTO, is used to determine the recovery strategy. A cold site cannot be used if the MTD is 12 h, for example. As a general rule, the shorter the MTD, the more expensive the recovery solution will be.

Redundant site

A *redundant site* is an exact production duplicate of a system that has the capability to seamlessly operate all necessary IT operations without loss of services to the end user of the system. A redundant site receives data backups in real time so that in the event of a disaster, the users of the system have no loss of data. It is a building configured exactly like the primary site and is the most expensive recovery option because it effectively more than doubles the cost of IT operations. To be fully redundant, a site must have real-time data backups to the redundant system and the end user should not notice any difference in IT services or operations in the event of a disruptive event.

Hot site

A *hot site* is a location that an organization may relocate to following a major disruption or disaster. It is a datacenter with a raised floor, power, utilities, computer peripherals, and fully configured computers. The hot site will have all necessary hardware and critical applications data mirrored in real time. A hot site will have the capability to allow the organization to resume critical operations within a very short period of time, sometimes in less than an hour.

It is important to note the difference between a hot site and a redundant site. Hot sites can quickly recover critical IT functionality; it may even be measured in minutes instead of hours. However, a redundant site will appear as operating normally to the end user no matter what the state of operations is for the IT program. A hot site has all the same physical, technical, and administrative controls implemented of the production site.

Warm site

A *warm site* has some aspects of a hot site; for example, readily accessible hardware and connectivity, but it will have to rely upon backup data in order to reconstitute a system after a disruption. It is a datacenter with a raised floor, power, utilities, computer peripherals, and fully configured computers.

Cold site

A *cold site* is the least expensive recovery solution to implement. It does not include backup copies of data, nor does it contain any immediately available hardware. After a disruptive event, a cold site will take the longest amount of time of all recovery solutions to implement and restore critical IT services for the organization. Especially in a disaster area, it could take weeks to get vendor hardware shipments in place, so organizations using a cold site recovery solution will have to be able to withstand a significantly long MTD measured in weeks, not days. A cold site is typically a datacenter with a raised floor, power, utilities, and physical security, but not much beyond that.

Reciprocal agreement

Reciprocal agreements are a bidirectional agreement between two organizations in which one organization promises another organization that it can move in and share space if it experiences a disaster. It is documented in the form of a contract written to gain support from outside organizations in the event of a disaster. They are also referred to as mutual aid agreements and they are structured so that each organization will assist the other in the event of an emergency.

Mobile site

Mobile sites are veritable datacenters on wheels in that they are towable trailers that contain racks of computer equipment, as well as HVAC, fire suppression, and physical security. They are a good fit for disasters such as a datacenter flood, where the datacenter is damaged but the rest of the facility and surrounding property are intact. They may be towed onsite, supplied with power and a network, and brought online.

RELATED PLANS

As discussed previously, the BCP is an umbrella plan that contains other plans. In addition to the DRP, other plans include the *COOP, the Business Resumption/Recovery Plan (BRP), Continuity of Support Plan, Cyberincident Response Plan, OEP, and the Crisis Management Plan (CMP)*. Table 7.3, from NIST Special Publication 800-34, summarizes these plans.

Table 7.3 Summary of BCP Plans From NIST SP 800-34[2]

Plan	Purpose	Scope
Business Continuity Plan (BCP)	Provide procedures for sustaining essential business operations while recovering from a significant disruption	Addresses business processes; IT addressed based only on its support for business process
Business Recovery (or Resumption) Plan (BRP)	Provide procedures for recovering business operations immediately following a disaster	Addresses business processes; not IT-focused; IT addressed based only on its support for business process
Continuity of Operations Plan (COOP)	Provide procedures and capabilities to sustain an organization's essential, strategic functions at an alternate site for up to 30 days	Addresses the subset of an organization's missions that are deemed most critical; usually written at headquarters level; not IT-focused
Continuity of Support Plan/IT Contingency Plan	Provide procedures and capabilities for recovering a major application or general support system	Same as IT contingency plan; addresses IT system disruptions; not business process-focused
Crisis Communications Plan	Provides procedures for disseminating status reports to personnel and the public	Addresses communications with personnel and the public; not IT-focused
Cyberincident Response Plan	Provide strategies to detect, respond to, and limit consequences of malicious cyber incident	Focuses on information security responses to incidents affecting systems and/or networks
Disaster Recovery Plan (DRP)	Provide detailed procedures to facilitate recovery of capabilities at an alternate site	Often IT-focused; limited to major disruptions with long-term effects
Occupant Emergency Plan (OEP)	Provide coordinated procedures for minimizing loss of life or injury and protecting property damage in response to a physical threat	Focuses on personnel and property particular to the specific facility; not business process or IT system functionality based

Continuity of operations plan

The COOP describes the procedures required to maintain operations during a disaster. This includes transfer of personnel to an alternate disaster recovery site and operations of that site.

Business recovery plan

The BRP, also known as the Business Resumption Plan, details the steps required to restore normal business operations after recovering from a disruptive event. This may include switching operations from an alternate site back to a repaired primary site.

The BRP picks up when the COOP is complete. This plan is narrow and focused: the BRP is sometimes included as an appendix to the BCP.

Continuity of support plan

The Continuity of Support Plan focuses narrowly on support of specific IT systems and applications. It is also called the IT Contingency Plan, emphasizing IT over general business support.

Cyberincident response plan

The Cyberincident Response Plan is designed to respond to disruptive cyberevents, including network-based attacks, worms, computer viruses, Trojan horses, etc. For example, self-propagating malicious code such as a worm has the potential to disrupt networks. Loss of network connectivity alone may constitute a disaster for many organizations.

Occupant emergency plan

The OEP provides the "response procedures for occupants of a facility in the event of a situation posing a potential threat to the health and safety of personnel, the environment, or property. Such events would include a fire, hurricane, criminal attack, or a medical emergency."[2] This plan is facilities-focused, as opposed to business- or IT-focused.

The OEP is focused on safety and evacuation, and should describe specific safety drills, including evacuation or fire drills. Specific safety roles should be described, including safety warden and meeting point leader, as described in Chapter 3.

Crisis management plan

The *CMP* is designed to provide effective coordination among the managers of the organization in the event of an emergency or disruptive event. The CMP details the actions management must take to ensure that life and safety of personnel and property are immediately protected in case of a disaster.

Crisis communications plan

A critical component of the CMP is the Crisis Communications Plan, which is sometimes simply called the communications plan. This is a plan for communicating to staff and the public in the event of a disruptive event. Instructions for notifying the affected members of the organization are an integral part to any BCP/DRP.

It is often said that bad news travels fast. Also, in the event of a postdisaster information vacuum, bad information will often fill the void. Public relations professionals understand this risk and know to consistently give the organization's "official story," even when there is little to say. All communication with the public should be channeled via senior management or the public relations team.

CALL TREES

A key tool leveraged for staff communication by the Crisis Communications Plan is the Call Tree, which is used to quickly communicate news throughout an organization without overburdening any specific person. The call tree works by assigning each employee a small number of other employees they are responsible for calling in an emergency event. For example, the organization's president may notify his board of directors of an emergency situation and they, in turn, will notify their top-tier managers. The top-tier managers will then call the people they have been assigned to call. The call tree continues until all affected personnel have been contacted.

The call tree is most effective when there is a two-way reporting of successful communication. For example, each member of the board of directors would report back to the president when each of their assigned call tree recipients had been contacted and had made contact with their subordinate personnel. Remember that cell phones and landlines may become congested or unusable during a disaster; the call tree should contain alternate contact methods in case the primary methods are unavailable.

EMERGENCY OPERATIONS CENTER

The Emergency Operations Center (EOC) is the command post established during or just after an emergency event. Placement of the EOC will depend on resources that are available. For larger organizations, the EOC may be a long distance away from the physical emergency; however, protection of life and personnel safety is always of the utmost importance.

BACKUPS AND AVAILABILITY

Though many organizations are diligent in going through the process of creating backups, verification of recoverability from those backup methods is at least as important and is often overlooked. When the detailed recovery process for a given backup solution is thoroughly reviewed, some specific requirements will become obvious. One of the most important points to make when discussing backup with respect to disaster recovery and business continuity is to ensure that critical backup media is stored offsite. Further, that offsite location should be situated such that, during a disaster event, the organization can efficiently access the media with the purpose of taking it to a primary or secondary recovery location.

HARDCOPY DATA

In the event that there is a disruptive event, such as a natural disaster that disables the local power grid, and power dependency is problematic, there is the potential to operate the organization's most critical functions using only hardcopy data. *Hardcopy data* is any data that are accessed through reading or writing on paper rather than processing through a computer system.

ELECTRONIC BACKUPS

Electronic backups are archives that are stored electronically and can be retrieved in case of a disruptive event or disaster. Choosing the correct data backup strategy is dependent upon how users store data, the availability of resources and connectivity, and what the ultimate recovery goal is for the organization.

Preventative restoration is a recommended control; an organization can restore data to test the validity of the backup process. If a reliable system, such as a mainframe, copies data to tape every day for years, what assurance does the organization have that the process is working? Do the tapes and the data they contain have integrity?

Full backups

A full system backup means that every piece of data is copied and stored on the backup repository. Conducting a full backup is time consuming and a strain on bandwidth and resources. However, full backups will ensure that any and all necessary data is protected.

Incremental backups

Incremental backups archive data that have changed since the last full or incremental backup. For example, a site performs a full backup every Sunday, with daily incremental backups from Monday through Saturday. If data is lost after the Wednesday incremental backup, four tapes are required for restoration: the Sunday full backup, as well as the Monday, Tuesday, and Wednesday incremental backups.

Differential backups

Differential backups operate in a similar manner as the incremental backups except for one key difference: differential backups archive data that have changed since the last full backup.

For example, the same site in our previous example switches to differential backups. They lose data after the Wednesday differential backup. Now only two tapes are required for restoration: the Sunday full backup and the Wednesday differential backup.

Tape rotation methods

A common tape rotation method is called *FIFO* (First In, First Out). Assume you are performing full daily backups and have 14 rewritable tapes total. FIFO (also called round robin) means you will use each tape in order and cycle back to the first tape

after the 14th is used. This ensures 14 days of data is archived. The downside of this plan is you only maintain 14 days of data; this schedule is not helpful if you seek to restore a file that was accidentally deleted 3 weeks ago.

Grandfather-Father-Son (GFS) addresses this problem. There are 3 sets of tapes: 7 daily tapes (the son), 4 weekly tapes (the father), and 12 monthly tapes (the grandfather). Once per week, a son tape graduates to father. Once every 5 weeks a father tape graduates to grandfather. After running for a year, this method ensures there are backup tapes available for the past 7 days, weekly tapes for the past 4 weeks, and monthly tapes for the past 12 months.

Electronic vaulting

Electronic vaulting is the batch process of electronically transmitting data that is to be backed up on a routine, regularly scheduled time interval. It is used to transfer bulk information to an offsite facility. There are a number of commercially available tools and services that can perform electronic vaulting for an organization. Electronic vaulting is a good tool for data that need to be backed up on a daily or possibly even hourly rate. It solves two problems at the same time: it stores sensitive data offsite and it can perform the backup at very short intervals to ensure that the most recent data is backed up.

Remote journaling

A database journal contains a log of all database transactions. Journals may be used to recover from a database failure. Assume a database checkpoint (snapshot) is saved every hour. If the database loses integrity 20 min after a checkpoint, it may be recovered by reverting to the checkpoint and then applying all subsequent transactions described by the database journal.

Remote journaling saves the database checkpoints and database journal to a remote site. In the event of failure at the primary site, the database may be recovered.

Database shadowing

Database shadowing uses two or more identical databases that are updated simultaneously. The shadow database(s) can exist locally, but it is best practice to host one shadow database offsite. The goal of database shadowing is to greatly reduce the recovery time for a database implementation. Database shadowing allows faster recovery when compared with remote journaling.

HA options

Increasingly, systems are being required to have effectively zero downtime, or an MTD of zero. Recovery of data on tape is certainly ill equipped to meet these availability demands. The immediate availability of alternate systems is required should a failure or disaster occur. A common way to achieve this level of uptime requirement is to employ a high availability cluster.

The goal of a high availability cluster is to decrease the recovery time of a system or network device so that the availability of the service is less affected than it would be by having to rebuild, reconfigure, or otherwise stand up a replacement system.

FAST FACTS

Two typical deployment approaches exist:

- *Active-active cluster* involves multiple systems, all of which are online and actively processing traffic or data. This configuration is also commonly referred to as load balancing and is especially common with public facing systems, such as Web server farms.
- *Active-passive cluster* involves devices or systems that are already in place, configured, powered on, and ready to begin processing network traffic should a failure occur on the primary system. Active-passive clusters are often designed such that any configuration changes made on the primary system or device are replicated to the standby system. Also, to expedite the recovery of the service, many failover cluster devices will automatically begin to process services on the secondary system should a disruption impact the primary device. It can also be referred to as a hot spare, standby, or failover cluster configuration.

DRP TESTING, TRAINING, AND AWARENESS

Testing, training, and awareness must be performed for the "disaster" portion of a BCP/DRP. Skipping these steps is one of the most common BCP/DRP mistakes. Some organizations "complete" their DRP, consider the matter resolved, and put the big DRP binder on a shelf to collect dust. This thought process is wrong on numerous levels.

First, a DRP is never complete but is rather a continually amended method for ensuring the ability for the organization to recover in an acceptable manner. Second, while well-meaning individuals carry out the creation and update of a DRP, even the most diligent of administrators will make mistakes. To find and correct these issues prior to their hindering recovery in an actual disaster, testing must be carried out on a regular basis. Third, any DRP that will be effective will have some inherent complex operations and maneuvers to be performed by administrators. There will always be unexpected occurrences during disasters, but each member of the DRP should be exceedingly familiar with the particulars of their role in a DRP, which is a call for training on the process.

Finally, it is important to be aware of the general user's role in the DRP, as well as the organization's emphasis on ensuring the safety of personnel and business operations in the event of a disaster. This section will provide details on steps to effectively test, train, and build awareness for the organization's DRP.

DRP TESTING

In order to ensure that a DRP represents a viable plan for recovery, thorough testing is needed. Given the DRP's detailed tactical subject matter, it should come as no surprise that routine infrastructure, hardware, software, and configuration changes will alter the way the DRP needs to be carried out. Organizations' information systems are in a constant state of flux, but unfortunately, much of these changes do not readily make their way into an updated DRP. To ensure both the initial and continued efficacy of the DRP as a feasible recovery methodology, testing needs to be performed.

DRP review

The DRP Review is the most basic form of initial DRP testing and is focused on simply reading the DRP in its entirety to ensure completeness of coverage. This review is typically performed by the team that developed the plan and will involve team members reading the plan in its entirety to quickly review the overall plan for any obvious flaws. The DRP Review is primarily just a sanity check to ensure that there are no glaring omissions in coverage or fundamental shortcomings in the approach.

Read-through

Read-through (also known as *checklist* or *consistency*) testing lists all necessary components required for successful recovery and ensures that they are or will be readily available should a disaster occur. For example, if the disaster recovery plan calls for the reconstitution of systems from tape backups at an alternate computing facility, the site in question should have an adequate number of tape drives on hand to carry out the recovery in the indicated window of time. The read-through test is often performed concurrently with the structured walkthrough or tabletop testing as a solid first-testing threshold. The read-through test is focused on ensuring that the organization has or can acquire in a timely fashion sufficient levels of resources upon which successful recovery is dependent.

Walkthrough/tabletop

Another test that is commonly completed at the same time as the checklist test is that of the *walkthrough*, which is also often referred to as a *structured walkthrough* or *tabletop exercise*. During this type of DRP test, which is usually performed prior to more in-depth testing, the goal is to allow individuals who are knowledgeable about the systems and services targeted for recovery to thoroughly review the overall approach. The term structured walkthrough is illustrative, as the group will discuss the proposed recovery procedures in a structured manner to determine whether there are any noticeable omissions, gaps, erroneous assumptions, or simply technical missteps that would hinder the recovery process from successfully occurring.

Simulation test/walkthrough drill

A *simulation test*, also called a *walkthrough drill* (not to be confused with the discussion-based structured walkthrough), goes beyond talking about the process and actually has teams to carry out the recovery process. A simulated disaster to which the team must respond as they are directed to by the DRP. As smaller disaster simulations are successfully managed, the scope of simulations will vary significantly and tend to grow more complicated and involve more systems.

Parallel processing

Another type of DRP test is *parallel processing*. This type of test is common in environments where transactional data is a key component of the critical business processing. Typically, this test will involve recovery of critical processing components at an alternate computing facility and then restore data from a previous backup. Note that regular production systems are not interrupted.

The transactions from the day after the backup are then run against the newly restored data, and the same results achieved during normal operations for the date in question should be mirrored by the recovery system's results. Organizations that are highly dependent upon mainframe and midrange systems will often employ this type of test.

Partial and complete business interruption

Arguably, the highest fidelity of all DRP tests involves *business interruption testing*. However, this type of test can actually be the cause of a disaster, so extreme caution should be exercised before attempting an actual interruption test. As the name implies, the business interruption style of testing will have the organization actually stop processing normal business at the primary location and will instead leverage the alternate computing facility. These types of tests are more common in organizations where fully redundant, often load-balanced operations already exist.

FAST FACTS

Each DRP testing method varies in complexity and cost, and simpler tests are less expensive. Here are the plans, ranked in order of cost and complexity, from low to high:

- DRP Review
- Read-Through/Checklist/Consistency
- Structured Walkthrough/Tabletop
- Simulation Test/Walkthrough Drill
- Parallel Processing
- Partial Interruption
- Complete Business Interruption

CONTINUED BCP/DRP MAINTENANCE

Once the initial BCP/DRP plan is completed, tested, trained, and implemented, it must be kept up to date. Business and IT systems change quickly, and IT professionals are accustomed to adapting to that change. BCP/DRP plans must keep pace with all critical business and IT changes.

CHANGE MANAGEMENT

Change management includes tracking and documenting all planned changes, including formal approval for substantial changes and documentation of the results of the completed change. All changes must be auditable.

CRUNCH TIME

The change control board manages this process. The BCP team should be a member of the change control board and attend all meetings. The goal of the BCP team's involvement on the change control board is to identify any changes that must be addressed by the BCP/DRP plan.

BCP/DRP MISTAKES

BCP and DRP are a business' last line of defense against failure. If other controls have failed, BCP/DRP is the final control. If it fails, the business may fail.

The success of BCP/DRP is critical, but many plans fail. The BCP team should consider the failure of other organizations' plans and view their own procedures under intense scrutiny. They should ask themselves this question: "Have we made mistakes that threaten the success of our plan?"

FAST FACTS

Common BCP/DRP mistakes include:

- Lack of management support
- Lack of business unit involvement
- Lack of prioritization among critical staff
- Improper (often overly narrow) scope
- Inadequate telecommunications management
- Inadequate supply chain management
- Incomplete or inadequate CMP
- Lack of testing
- Lack of training and awareness
- Failure to keep the BCP/DRP plan up to date

SPECIFIC BCP/DRP FRAMEWORKS

Given the patchwork of overlapping terms and processes used by various BCP/DRP frameworks, this chapter focused on universal best practices without attempting to map to a number of different (and sometimes inconsistent) terms and processes described by various BCP/DRP frameworks.

A handful of specific frameworks are worth discussing, including NIST SP 800-34, ISO/IEC-27031, and BCI.

NIST SP 800-34

The National Institute of Standards and Technology (NIST) Special Publication 800-34 Rev. 1 "Contingency Planning Guide for Federal Information Systems" may be downloaded at http://csrc.nist.gov/publications/nistpubs/800-34-rev1/sp800-34-rev1_errata-Nov11-2010.pdf. The document is of high quality and is in the public domain. Plans can sometimes be significantly improved by referencing SP 800-34 when writing or updating a BCP/DRP.

ISO/IEC-27031

ISO/IEC-27031 is a new guideline that is part of the ISO 27000 series, which also includes ISO 27001 and ISO 27002 (discussed in Chapter 2). ISO/IEC 27031 focuses on BCP (DRP is handled by another framework; see below).

FAST FACTS

According to http://www.iso27001security.com/html/27031.html, ISO/IEC 27031 is designed to:

- "Provide a framework (methods and processes) for any organization—private, governmental, and nongovernmental
- Identify and specify all relevant aspects including performance criteria, design, and implementation details for improving ICT readiness as part of the organization's ISMS, helping to ensure business continuity
- Enable an organization to measure its continuity, security and hence readiness to survive a disaster in a consistent and recognized manner."[4]

Terms and acronyms used by ISO/IEC 27031 include:

- ICT—Information and Communications Technology
- ISMS—Information Security Management System

A separate ISO plan for disaster recovery is ISO/IEC 24762:2008, "Information technology—Security techniques—Guidelines for information and communications technology disaster recovery services." More information is available at http://www.iso.org/iso/catalogue_detail.htm?csnumber=41532

BS-25999 AND ISO 22301

British Standards Institution (BSI, http://www.bsigroup.co.uk/) released BS-25999, which is in two parts:

- "Part 1, the Code of Practice, provides business continuity management best practice recommendations. Please note that this is a guidance document only.
- Part 2, the Specification, provides the requirements for a Business Continuity Management System (BCMS) based on BCM best practice. This is the part of the standard that you can use to demonstrate compliance via an auditing and certification process."[5]

BS-25999-2 has been replaced with ISO 22301:2012 Societal security—Business continuity management systems—Requirements. "ISO 22301 will supersede the original British standard, BS 25999-2 and builds on the success and fundamentals of this standard. BS ISO 22301 specifies the requirements for setting up and managing an effective BCMS for any organization, regardless of type or size. BSI recommends that every business has a system in place to avoid excessive downtime and reduced productivity in the event of an interruption."[6]

BCI

The Business Continuity Institute (BCI, http://www.thebci.org/) published a six-step Good Practice Guidelines (GPG), most recently updated in 2013: "The Good Practice Guidelines (GPG) are the independent body of knowledge for good Business Continuity practice worldwide. They represent current global thinking in good Business

Continuity (BC) practice and now include terminology from ISO 22301:2012, the International Standard for Business Continuity management systems."[7]

FAST FACTS

GPG 2013 describes six Professional Practices (PP).

- Management Practices
  - PP1 Policy and Program Management
  - PP2 Embedding Business Continuity
- Technical Practices
  - PP3 Analysis
  - PP4 Design
  - PP5 Implementation
  - PP6 Validation[8]

SUMMARY OF EXAM OBJECTIVES

In this chapter, we have discussed operational security. Operations security concerns the security of systems and data while being actively used in a production environment. Ultimately, operations security is about people, data, media, and hardware, all of which are elements that need to be considered from a security perspective. The best technical security infrastructure in the world will be rendered moot if an individual with privileged access decides to turn against the organization and there are no preventive or detective controls in place within the organization.

We also discussed Business Continuity and Disaster Recovery Planning, which serve as an organization's last control to prevent failure. Of all controls, a failed BCP or DRP can be most devastating, potentially resulting in organizational failure, injury, or loss of life.

TOP FIVE TOUGHEST QUESTIONS

1. Which plan details the steps required to restore normal business operations after recovering from a disruptive event?
   - **A.** Business Continuity Plan (BCP)
   - **B.** Business Resumption Plan (BRP)
   - **C.** Continuity of Operations Plan (COOP)
   - **D.** Occupant Emergency Plan (OEP)
2. What metric describes how long it will take to recover a failed system?
   - **A.** Minimum Operating Requirements (MOR)
   - **B.** Mean Time Between Failures (MTBF)
   - **C.** The Mean Time to Repair (MTTR)
   - **D.** Recovery Point Objective (RPO)

3. What metric describes the moment in time in which data must be recovered and made available to users in order to resume business operations?
   A. Mean Time Between Failures (MTBF)
   B. The Mean Time to Repair (MTTR)
   C. Recovery Point Objective (RPO)
   D. Recovery Time Objective (RTO)
4. Maximum Tolerable Downtime (MTD) is comprised of which two metrics?
   A. Recovery Point Objective (RPO) and Work Recovery Time (WRT)
   B. Recovery Point Objective (RPO) and Mean Time to Repair (MTTR)
   C. Recovery Time Objective (RTO) and Work Recovery Time (WRT)
   D. Recovery Time Objective (RTO) and Mean Time to Repair (MTTR)
5. Which level of RAID does NOT provide additional reliability?
   A. RAID 1
   B. RAID 5
   C. RAID 0
   D. RAID 3

ANSWERS

1. Correct answer and explanation: B. Business Resumption Planning details the steps required to restore normal business operations after a recovering from a disruptive event.
   Incorrect answers and explanations: Answers A, C, and D are incorrect. Business Continuity Planning develops a long-term plan to ensure the continuity of business operations. The Continuity of Operations Plan describes the procedures required to maintain operations during a disaster. The Occupant Emergency Plan provides the response procedures for occupants of a facility in the event a situation poses a threat to the health and safety of personnel, the environment, or property.
2. Correct answer and explanation: C. The Mean Time to Repair (MTTR) describes how long it will take to recover a failed system. It is the best estimate for reconstituting the IT system so that business continuity may occur.
   Incorrect answers and explanations: A, B, and D. Answers A, B, and D are incorrect. Minimum Operating Requirements describe the minimum environmental and connectivity requirements in order to operate computer equipment. Mean Time Between Failures quantifies how long a new or repaired system will run before failing. The Recovery Point Objective (RPO) is the moment in time in which data must be recovered and made available to users in order to resume business operations.
3. Correct Answer and Explanation: C. The Recovery Point Objective (RPO) is the moment in time in which data must be recovered and made available to users in order to resume business operations.

Incorrect answers and explanations: Answers A, B, and D are incorrect. Mean Time Between Failures quantifies how long a new or repaired system will run before failing. Mean Time to Repair describes how long it will take to recover a failed system. Recovery Time Objective describes the maximum time allowed to recover business or IT systems.

4. Correct answer and explanation: C. The Recovery Time Objective (RTO, the time it takes bring a failed system back online) and Work Recovery Time (WRT, the time required to configure a failed system) are used to calculate the Maximum Tolerable Downtime. RTO + WRT = MTD.
   Incorrect answers and explanations: Answers A, B, and D are incorrect. Maximum Tolerable Downtime does not directly use Recovery Point Objective or Mean Time to Repair as metrics.
5. Correct answer and explanation: C. RAID 0 provides only striping and is used simply for performance purposes. It offers no additional data redundancy or resiliency.
   Incorrect answers and explanations: Answers A, B, and D are incorrect. RAIDs 1, 3, and 5 all provide reliability gains through either mirroring or parity measures.

ENDNOTES

1. *NIST special publication 800–61: computer security incident handling guide*. http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf [accessed 26.04.16].
2. Swanson M, Wohl A, Pope L, Grance T, Hash J, Thomas R. *NIST SP 800-34 contingency planning guide for information technology systems*. https://www.fismacenter.com/sp800-34.pdf [accessed 26.04.16].
3. *Understanding security risk management: recovery time requirements*. http://searchsecuritychannel.techtarget.com/generic/0,295582,sid97_gci1268749,00.html [accessed 26.04.16].
4. *ISO/IEC 27031:2011 Information technology—security techniques—guidelines for information and communications technology readiness for business continuity*. http://www.iso27001security.com/html/27031.html [accessed 26.04.16].
5. *ISO 22301 business continuity standard in IT*. http://eradar.eu/business-continuity/ [accessed 26.04.16].
6. *Moving from BS 25999-2 to ISO 22301*. http://www.bsigroup.com/Documents/iso-22301/resources/BSI-BS25999-to-ISO22301-Transition-UK-EN.pdf [accessed 26.04.16].
7. *The good practice guidelines*. http://www.thebci.org/index.php/resources/the-good-practice-guidelines [accessed 26.04.16].
8. *Good practice guidelines 2013 global edition*. http://www.thebci.org/index.php/the-gpg-lite [accessed 26.04.16].